IN THE CLAIMS

The listing of claims will replace all prior versions, and listings, of claims in the application.

1. (Currently Amended) A method for delivery of high-performance online multimedia services comprising:

assigning general content to be multicast;

customizing the general content to suit a first area, which corresponds to a first geographical region, and thus forming a first version of the content;

customizing the general content to suit a second area, which corresponds to a second geographical region, and thus forming a second version of the content;

multicasting, thorugh a first head-end, the first version to an end-user system in the first area;

caching the first version of the content at the first head-end; [[and]]

multicasting, through a second head-end, the second version to an end-user system in the second area;

caching the second version of the content at the second head-end;

receiving a request for the first version of the content at the first head-end from a second end-user system in the first area; and

providing the first version of the content from cache in response to the request.

2-3. (Canceled)

4. (New) The method of claim 1, further comprising:

receiving, at a first regional server in communication with the first head-end and a third head-end, the the first version of the content from the remote source and en route to the first head-end;

caching the first version of the conten at the first regional server; and

receiving the request for the first version of the content from a third end-user system in communication with the third head-end; and

providing the first version of the content from the first regional server.

5. (New) The method of claim 4, further comprising:

receiving a request for uncached content; and

bypassing a backbone network connection by directly connecting to a content source.

6. (New) A computer-readable medium storing a computer program product configured to perform a method for delivery of high-performance online multimedia services, the method comprising:

assigning general content to be multicast;

customizing the general content to suit a first area, which corresponds to a first geographical region, and thus forming a first version of the content;

customizing the general content to suit a second area, which corresponds to a second geographical region, and thus forming a second version of the content;

multicasting, thorugh a first head-end, the first version to an end-user system in the first area;

caching the first version of the content at the first head-end;

multicasting, through a second head-end, the second version to an end-user system in the second area;

caching the second version of the content at the second head-end;

receiving a request for the first version of the content at the first head-end from a second end-user system in the first area; and

providing the first version of the content from cache in response to the request.

7. (New) The computer program product of claim 6, the method further comprising:

receiving, at a first regional server in communication with the first head-end and a third head-end, the the first version of the content from the remote source and en route to the first head-end;

caching the first version of the conten at the first regional server; and

receiving the request for the first version of the content from a third end-user system in communication with the third head-end; and

providing the first version of the content from the first regional server.

8. (New) The method of claim 7, further comprising:

receiving a request for uncached content; and

bypassing a backbone network connection by directly connecting to a content source for requesting the uncached content.

9. (New) A system for delivery of high-performance online multimedia services, the method comprising:

a content source to assign general content to be multicast, to customize the general content to suit a first area, which corresponds to a first geographical region, and thus forming a first version of the content, to customize the general content to suit a second area, which corresponds to a second geographical region, and thus forming a second version of the content, and to multicast the first version to an end-user system in the first area;

a first head-end, in communication with the content source, to cache the first version of the content; and

a second head-end, in communication with the content source, to cache the second version of the content,

wherein the first head-end receives a request for the first version of the from a second end-user system in the first area, and provides the first version of the content from cache.

10. (New) The system of claim 9, further comprising:

a regional server, in communication with the content source, and also in communication with the first head-end and a third head-end, to receive the the first version of the content from the remote source and en route to the

first head-end, and to cache the first version of the content at the first regional server, wherein the regional server receives a request for the first version of the content from a third end-user system in communication with the third head-end, and provides the first version of the content from cache.

11. (New) The system of claim 10, wherein the regional server, in response to receiving a request for uncached content, bypasses a backbone network connection to directly connect to the content source for requesting the uncached content.